ChinaXiv合作期刊 图书情报工作
第65卷 第4期 2021年2月

突破性创新早期识别与弱信号分析综述*

■ 刘亚辉[1,2]　许海云[3,4]

[1] 中国科学院成都文献情报中心　成都 610041　2 中国科学院大学经济与管理学院图书情报与档案管理系　北京 100190

[3] 山东理工大学管理学院　淄博 255000　[4] 中国科学技术信息研究所　北京 100038

摘　要：[目的/意义]通过比较分析不同的突破性创新识别方法，总结现有方法存在的问题，将弱信号引入突破性创新研究的识别中，重点关注突破性创新早期的各类弱信号，尤其是弱关联关系分析，以期实现早期预判。[方法/过程]首先，通过调研现有的识别方法，提炼当前存在的主要问题，指出研究弱信号的必要性。之后，从不同的学科角度介绍弱信号的内涵及表征形式，对其特征进行概括，梳理弱信号的几种识别方法。最后介绍弱关系分析的内涵及应用，提出借鉴多元关系融合算法模型可以实现多种弱关系的有效融合，获取更明确的信息。[结果/结论]突破性创新的识别研究中受关注最多是文献间的引用关系、主题词之间的语义关系等强关系数据，而弱关系蕴含着更多元化的信息，加强弱关系分析可以实现对突破性创新研究的早期预判。未来研究需要寻求有效捕捉弱关联的方法，注重主题的动态演化规律，如利用高阶网络模型分析有效弱信号，提高突破性创新早期识别的准确性。

关键词：突破性创新　早期识别　弱信号　弱关系　关系融合

分类号：G250.2

DOI：10.13266/j.issn.0252-3116.2021.04.010

1 引言

当今世界正孕育着第六次科技革命，科学技术体系不断涌现重大突破，并深刻改变着社会的经济面貌。相对于渐进性研究，突破性创新的创新程度更高，代表着科技创新过程中最具前瞻性的发展方向。如何从新兴主题中识别和预判具有突破性创新的重大发现，对政策的制定和企业的战略布局都至关重要。尽可能早地识别出重大突破，政府和企业就能够及时调整科技战略，合理分配研发资源，有更多的应对时间抢得发展先机。但是许多科学突破和创造性的发现，在发展早期都具有较高的不确定性和模糊性，增加了早期迹象分析的难度。因此，如何在研究项目的开始时期关注早期迹象，预判其未来的变革潜力是科学政策和研究评估中亟待解决与最具挑战性的问题。

尽管已经存在着多种突破性研究的识别方法，但识别结果大多已是在研究领域的快速发展期或成熟期，难以做到早期发现。目前以专家判断为主的定性方法是研判未来科技发展趋势的重要手段，但是随着学科融合与数据的爆炸式涌现，专家智慧难以快速、准确的发挥最佳效能。探测突破性创新常用的定量方法，如引文网络分析、主题突变检测、睡美人文献分析及技术演化法，多注重高引用、强共现等"强信号"分析，尚未深入利用有助于早期识别的"弱信号"。弱信号预示着未来的变化，具体表征形式多是某种现象、事件、机会或威胁，分析弱信号可以辅助更早地了解和控制事物的未来发展。

鉴于此，本研究遵循科学发展规律，考虑科学发展过程中的不确定性，综述弱信号分析方法及将其应用于突破性创新识别的可能性。本文首先介绍突破性创新的相关概念，总结突破性创新的非线性、不确定性等特征；之后梳理现有突破性研究的识别方法，总结现有方法存在的问题；在此基础上引入弱信号，介绍其概念、特征及识别方法，阐述弱信号用于突破性创新早期

* 本文系国家自然科学基金项目"基于科学-技术主题关联分析的创新演化路径识别方法研究"（项目编号：71704170）、国家重点研发计划项目课题"颠覆性技术地平线扫描系统"（项目编号：2019YFA0707202-01）和国家自然科学基金项目"科技关联视角下新兴技术弱信号扫描预判方法研究"（项目编号：72074014）"研究成果之一。

作者简介：刘亚辉（ORCID：0000-0002-3577-6111），硕士研究生；许海云（ORCID：0000-0002-7453-3331），教授，博士，通讯作者，E-mail：xuhaiyunnemo@gmail.com。

收稿日期：2020-06-27　修回日期：2020-09-29　本文起止页码：89-101　本文责任编辑：王传清

识别的可行性;最后给出未来突破性创新的早期识别研究中可以深入探讨的地方。

2 研究现状

2.1 突破性创新的内涵

2.1.1 突破性创新的定义

突破性创新起源于熊彼特的"创造性的破坏"[1],之后 G. Dosi 和 R. R. Nelson 发表的《技术范式与技术轨道》将突破性技术创新与渐进性技术创新统一到一个理论框架内,开启了对突破性创新的研究探索[2]。研究人员对于突破性创新的理解与认知不尽相同,通过调研发现,大多数研究的关注点集中在:①技术改进、产品性能的提高、市场价值;②科学研究中的突破性发现、全新的知识理论基础。如表1所示:

表1 突破性创新的相关定义

研究重点	概念内涵	代表性研究
技术改进、产品性能及市场价值提升	是一种非连续性的、革命性的创新,会有全新的科技知识和资源,并会淘汰现有的技术和产品	M. L. Tushman 和 P. Anderson[3]
	使产品、工艺或服务具有全新的性能特征或虽具有相似的特征但性能大幅提高且成本降低,或创造出能够改变现有市场和产业的新产品	V. Kotelnikov[4] 肖海林和董慈慈[5]
	是将新奇、独特和精妙的技术引入新的产品,并可以改变市场消费模式的一种创新,能显著提升技术水平和顾客价值	K. Z. Zhou[6] 孙晓雅和陈娟娟[7]
	是一种技术、产品、过程和商业模式上的创新,能对现有的产品和服务进行改造,赢得顾客市场	B. Fores 和 C. Camison[8]
	两个层次:一方面是现有技术的应用和组合产生的市场突破性;另一方面则是指技术层面的不连续性	张金柱和张晓林[9]
科学研究中的突破性创新	突破性创新常与科学革命联系在一起,使科学范式发生改变。科学发展是在常规研究与革命性研究两种状态的相互交替转换中进行的	T. S. Kuhn[10]
	突破性创新研究可以消除科学进步的障碍,可能创造新的理论或改进现有的理论,更好地描述已知的现象,但都受到科学规范的约束	K. B. Wray[11]
	突破性创新可能会彻底改变该领域的认知结构,延伸出新领域的研究,往往与当前的科学领域内占主导地位的理论框架不相容	H. Andersen 等[12] 杜建等[13]
	促进科学进展的大多数突破性创新都是较小的改进,很少有开创性的创新,但都是革命性的,驳斥了库恩常规科学与变革性科学交替出现的观点	P. Galison[14] A. V. Raan[15]
	突破性创新是一个过程,通过无数连续的微小的、渐进的进步产生一种思考问题的新方式,科学发展并没有界限分明的范式转化	J. R. Hollingsworth[16]

2.1.2 突破性创新的特征

突破性创新的特征较为复杂,在研究学界尚未形成一致和全面的解读。本文重点分析了研究人员对基础科学研究中突破性创新的特征描述,如表2所示:

表2 突破性创新的特征

特征	特征描述
新颖性/前瞻性[17-18]	突破性研究通常建立在新的科学原理或者技术手段之上,引领科技未来发展
非线性/不连续性[10,19-20]	科学发展是非线性的,突破性创新研究可能会引起不连续的断层现象
不确定性[1,21-22]	每一项突破都有偶然性的因素,突破性创新发展过程中存在很强的不确定性,预测难度较大

通过对突破性创新的定义及特征梳理可知,突破性创新对未来科技发展产生重大影响。科学发现与它们在技术上的结合,可能需要跨越数十年,在研究与开发(R&D)系统中通常以复杂的方式相互联系。所以本研究关注基础研究,聚焦于知识创新带来的科学进程中的突破性发现,暂不考虑技术领域以及对市场的不确定性影响。

2.1.3 突破性创新的分类及可预测性

突破性创新的产生契机及变革程度存在差别,因此突破性创新存在多种形式,其识别预测的难度也不同。当前还没有统一的分类标准,D. E. Koshland 回顾科学发展史将突破性创新分为3种类型,形成了有较大影响的"Cha-Cha-Cha"理论[23],如表3所示:

表3 突破性创新分类

类型	具体描述	侧重点
Charge 型	这类发现是针对大家每个人都会看到的现象,如苹果下坠、星星转动,思考这些现象的原因,给出理论来解释这些现象	现象已经存在,不需要刻意感知,重点在于揭示现象背后的原因
Challenge 型	"解释异常现象",即为当时的科学理论无法解释的事实提供新的概念、理论或新的解决方案	要求科研人员能够感知异常,意识到它们的重要性,设计新的概念或理论解释异常
Chance 型	被称为偶然发现,是研究人员偶然发现的现象或事实	主要是看到偶然的现象并研究偶然"事故"是如何引发的

从表3的分类结果可以看出,Chance 型科学突破是在特殊的环境或经历机缘巧合的事件被发现,除此

之外，大多数推动科学进程发展的突破性创新并不是偶然，需要研究人员思考各类事实和问题，将这些独立的、看似不相关的信息联系起来。

一项突破性创新的出现在现有的科学体系中常被认为是“不连续性”的，属于“间断性发展”。实际上如果“放大”断点之处的具体内容，从突破本身相关的内容或解决的问题来看，这种突破与先前的研究是存在关联的，它的产生、发展过程是“连续性”的，这是突破性创新可以预测的前提。本文在直角坐标系中绘制科学创新发展过程进行解释，图 1 中 O 为坐标原点、横轴为时间维度、纵轴为创新性维度。Ⅰ、Ⅱ、Ⅲ、Ⅳ为科学发展的 4 个渐进性创新阶段，A、B、C 表示 4 阶段过程中的“跳跃性”突破性创新。将其“放大”后可知，突破性创新的产生、发展是有迹可循的，有进行提前预测的可能性。

因此在“Cha-Cha-Cha”理论中，Chance 型这类偶然发现的重大突破性成果存在随机、不确定性因素，难以进行提前预测。而对于 Charge 型和 Challenge 型而言，大多数应该都是有迹可循的，研究通过深入挖掘突破性研究发表初期的“预警信号”，分析其背后的内在关联，为突破性创新的早期识别提供新思路。

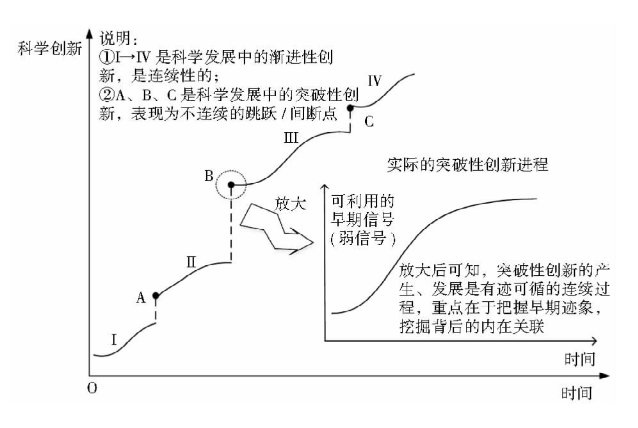


图 1　科学发展进程中的突破性创新

2.2　突破性创新识别方法

突破性创新的早期信号较少，识别难度较大。因为突破性创新往往与现有的科学理论相抵触，违背或超越了人们的认知和技术水平，科学共同体习惯进行抵制[24]。研究人员对突破性研究的主题识别方法进行了探索，以专家判断为主的定性分析方法依然是探测突破性研究的重要手段，但是在学科交叉融合的大科学时代，完全依靠专家个人对某个学科领域的了解和经验可能无法获得最准确的结果。数据处理、分析工具及算法层出不穷，成为专家判断过程中的有效辅助手段。通过梳理当前定量识别方法，可以发现不同的识别方法具有一定的适用范围，各有优缺点，如表 4 所示：

表 4　突破性研究主题识别方法对比分析

识别方法	方法说明	优势	不足	代表性研究
专家判断法	利用专家的智慧和经验识别预测科技发展态势	专家有完备的知识体系和充足的经验，适用于单一学科领域	主观性强；学科交叉融合的时代背景下，效率低，准确性不强	《麻省理工科技评论》[25]、《2019 科学发展报告》[26]
引文网络分析法	针对文献间的引用和被引用关系，构建网络，进行共被引网络分析或耦合分析	引文间体现了文献内容的相关性，可以客观反映科学领域的发展及学科间的关系	存在滞后性，只有当文献达到一定的被引次数才能得到关注	H. G. Small[27]、J. W. Schneider 和 R. Costas[28]、J. Wang 等[29]、J. Winnink[21]
主题突变监测法	关注主题词的变化及离群数据，通过分析突发词状态的变化发现突破性创新	研究对象为主题词，分析粒度更细，动态性和时效性更强	忽视文献间和主题词间的语义关联，不能很好地揭示学科领域的整体知识结构	李勇等[30]、J. Kleinberg[31]、张金柱和张晓林[32]、J. Yoon 和 K. Kim[33]、罗素平等[34]
睡美人文献分析法	对睡美人论文和专利深入分析，探索此类文献中可能孕育的潜在突破性研究	科学和技术领域都有延迟承认现象，缩短突破性创新成果的认可时滞	睡美人文献仅仅是科学文献中的一小部分，只能作为识别科学突破的补充手段	N. Palomeras[35]、杜建[36]、A. F. J. Van Raan 和 J. J. Winnink[37]
技术演化视角分析法	从技术轨道的转换论角度出发，在新的技术轨道未成型时，判断它的转换期	突破性创新常出现在新旧技术轨道的更迭期，该视角识别更准确，实施较为简便	不同学科领域内突破性创新的发展模式不同，难以提炼出普适性的衡量指标	A. Sood 和 G. J. Tellis[38]、G. J. TELLIS[39]、杨国忠和陈佳[40]
机器学习算法模型识别法	针对多源数据，借助神经网络及深度学习等算法模型，预测主题发展趋势	可以处理大体量的复杂数据，利用算法对参数模型进行优化，提高识别的准确度	需要多次试验寻找合适的算法模型，对研究人员的专业知识有一定要求	J. Mao 等[41]、徐路路和王芳[42]、白如江等[43]

2.3　识别方法存在的主要问题

2.3.1　侧重热点监测而非早期识别

从识别的阶段来看，当前研究并未有效捕捉突破性创新主题的“早期信号”，造成识别时滞。现有方法大多是在研究领域的快速发展期或成熟期进行热点主题的探测，而非早期识别，导致前瞻性预判不足[44]。很多突破性创新几乎没有早期迹象可供测度，阶段越早，信息越少，识别难度越大。知识的新颖度很难用可

量化的指标来计算,同时在创新过程中伴随着诸多不确定性,释放出的早期信号携带的信息也具有不确定性,而且对未来的影响也具有不确定性,因此难以在较早阶段探测突破性创新。

2.3.2　弱关系挖掘分析不够深入

在识别突破性创新的实现手段方面,现有方法大多针对“强信号”,如高被引、高下载量的文献开展分析。在主题聚类过程中,共词分析常采用的也是高频词,低频词所形成的主题簇团无法体现,不能很好地展现学科领域的发展全貌。突破性创新在萌芽之初,往往以主题间微弱关联(弱关系)的形式存在,这些弱关系代表了与高频词和强关系不同的数据对象,可能反映事物真实的性质,对该类型数据的忽略可能导致重要信息的遗漏[45]。但由于技术水平和研究人员认知水平的限制,弱关联关系仅作为一个附加关注对象,充当简单的辅助监测手段,而没有被充分挖掘,更缺乏能够明确地从弱信号视角识别突破性创新的研究。

2.3.3　缺乏普适的突破性创新主题演化规律

从整体的学科研究角度来看,识别突破性创新,需要了解其产生前后的知识状态,对知识扩散过程有整体性认识。主题演化规律可以揭示科技创新过程中主题的产生、发展、演变、消亡等过程,依据该规律通过分析某一学科领域现在的知识状态推断后续的发展趋势[46]。已有研究人员注意到这个问题,引入时间序列,关注主题词之间的深层次语义关联,尝试对主题的演化路径进行分析,捕获演化规律。但是结果往往具有局限性,主要是两方面的原因:一方面是科学发展受到很多不确定性因素的影响,难以用确定的规律来描绘路径;另一方面是设计的量化测量方法难以展现准确的主题演化规律,不具有领域的可迁移性,规律的有效性难以有效验证。

3　弱信号及其识别与应用

科学研究的本质在于探索未知,客观事物具有无限性、发展性,科学是一个不断修正的过程,研究过程受到很多不确定性因素的影响,因此科学结论都有它的条件和适用范围[47]。本文考虑科学发现中伴随的不确定性,将弱信号引入突破性研究的识别中。重点关注突破性创新的早期“预警信号”,借鉴各种弱信号捕获分析方法,解析零散弱信号之间的关联并加以深度挖掘,为突破性研究的早期预测提供可能性。

3.1　弱信号的相关概念

“信号”一词最早出现在电子通讯及军事领域,一般包含光信号、声信号和电信号。它是一种可以觉察的物理量或脉冲(如电压、电流、磁场强度等等),是运载信息的载体[48]。初期信号往往携带着预示未来发展的有效信息,但强度较弱,不易被感知与接收。目前,随着技术手段的不断发展,被噪声掩盖的微弱信号的检测越来越受到人们的重视,对弱信号的研究也渐渐由工程技术领域转向社会管理领域。通过调研相关文献,发现以下 3 个领域对弱信号的概念进行了诸多讨论:

3.1.1　市场经济学角度

市场经济学领域认为信号是帮助理性人做出价值判断的信息集合,它并不全是有形的,这一概念超越了通讯领域中信号的含义。美国经济学家斯宾塞首次运用“市场信号”这一概念,将其解释为市场上卖方向买方发出代表着产品或其他交易对象质量的信号,可减少信息不对称造成的影响[49]。O. Heil 和 T. S. Robertson 将市场信号定义为对潜在行动的宣告和预演,将市场信号与市场行为区别开来,认为市场信号仅能预示某种可能的市场行为,并不一定会真实发生,依据携带信息量大小可以分为强信号和弱信号[50]。G. S. Day 和 P. J. H. Schoemaker 指出弱信号是较为模糊的、不确定的,常与大量的噪音混合在一起,忽视这些信号可能会使企业丧失新的市场,将发展机会让给对手[51]。

3.1.2　企业战略规划角度

在企业战略规划方面,H. Ansoff 首先提出弱信号的概念,用于应对企业战略管理领域的战略突袭问题。他将弱信号描述为对企业未来发展具有不确切影响的早期迹象、征兆或事件,并将信号的强度和公司的分级响应联系起来[52-53]。B. Coffman 做了更详细的解释,提出弱信号存在的具体表征类型:可能影响企业环境发展的想法一个组织所面临的威胁或机会;企业成长和发展可以借助的事物等[54]。同样地,J. S. Brown 认为要获得并保持竞争优势,处于动荡环境中的公司需要扫描业务环境,以捕捉早期的微弱信号,经分析利用使之成为新的战略洞察力的来源[55]。赵小康将弱信号应用在竞争环境中,指出利用好弱信号,可以对产品发展变化、组织长期运行或行业趋势波动做出早期判断[56]。邓胜利等从竞争情报角度分析,认为弱信号是有预见性的、模糊零碎的、形式和来源多样的迹象符号,对企业面向未来的决策具有战略意义[57]。

3.1.3　未来学角度

以芬兰学者为首的研究人员从未来学的角度对弱信号的内涵进行了讨论。O. Kuusi 等运用德尔菲法协

调专家进行交流，整理出两种相互矛盾的观点：①弱信号是预示未来变化的早期预警，多个弱信号结合后信号增强。重要性取决于信号接收者的认知能力，通常被行业先驱先感知，而不是专家；②弱信号通常对未来结果产生根本影响，是一种客观现象，不依赖于接收者而存在。随着时间流逝逐渐增强，专家最有能力对其识别[58]。针对上述专家讨论结果中弱信号内涵的非一致性，M. Moijanen 总结了研究人员对弱信号的 3 种理解：弱信号既与现象相关联，也与结果相关；弱信号仅代表着正在变化的现象；弱信号只是一个预示着未来变化的征兆[59]。R. Pitkänen 反对在未来学的研究中使用“弱信号”这一术语，认为信号需要发送者，但弱信号的发送器往往不明确。未来学家对弱信号的讨论多为主观推断，缺乏严谨的理论定律[60]。E. Hiltunen 在符号三元组模型基础上，从信号的数量、信号背后蕴含的信息及对于信号的理解 3 个维度解读弱信号，并指出任一维度的上升都会使得弱信号增强[61]。

由各领域研究人员对于弱信号的研究可以看出，分析弱信号可以辅助尽早了解和控制事物的未来发展，在科技信息领域，也应该对不同表征形式的弱信号加以捕获、利用，尽早挖掘出其潜在价值。

3.2 弱信号的特征

研究人员对弱信号的内涵表述形式多样，但包含了诸多共同点，如影响具有不确定性、碎片化、解读的复杂性、可演化成趋势和引导未来变化等。本文归纳整理了弱信号的特征描述如表 5 所示：

表 5 弱信号的特征

特征	特征描述	代表性研究
碎片化	弱信号是不完全信息，广泛而分散，量小且难以理解，相关性不强，很少引起重视	H. Lesca 和 N. Lesca[62]、P. J. H. Schoemaker 等[63]、邓胜利等[57]
持续性	弱信号具有一定的持续性，它从第一次出现到成为强信号或消失需要持续一段时间	O. Kuusi 等[58]、单彬[64]
不确定性	表现在两个方面：弱信号蕴含的信息本身模糊不清；含义清晰明确，但对未来的影响具有不确定性	E. Hiltunen[65]、党倩娜[66]
前瞻预见性	弱信号只与某些潜在的事情或现象有关，其作用无法立即显现，但会带来较大的后果和影响	M. Moijanen[59]、P. Rossel[67]、C. Park 和 S. Cho[68]
主观性	弱信号代表的信息是客观存在的，但其效用和意义往往跟接收者紧密联系，这个层面来讲具有主观性	P. Saul[69]、T. Kuosa[70]、董尹等[71]

由表 5 可知，弱信号是突破性创新的早期表现形式。突破性创新在早期阶段是以弱信号的形式出现，随着时间的推移，弱信号逐渐增强，更容易被监测人员捕获。因此关注早期弱信号的获取与分析使得突破性创新的早期识别成为可能。

弱信号存在多种形式，B. Coffman 作出了很详细地划分，认为日常生活中有 3 种类型的弱信号在我们身边传递：已经超越了我们的感知，且没有能力接收；在我们的感知范围内，但就我们自身的认知能力无法识别；能够被识别并且会对我们的行为产生影响、带来改变[54]。因此，突破性创新在早期阶段如何捕获蕴含潜在价值的弱信号是研究的关键。

3.3 弱信号的识别方法

通常情况下，随着事物的产生、发展，有一些线索和迹象不断显露出来，经过积累、序化，形成具有明确指向性的信息。在整个过程中，弱信号由弱变强的过程，时间跨度较长[64]，要实现在更早阶段识别出弱信号并加以利用，应需要研究人员主动地发掘、筛选。弱信号最初在电子通信工程领域的分析利用较为广泛，研究人员常运用混沌振子检测法探测弱信号。C. Deng 和 S. Zhang 将 Lorenz 系统方法与遗传算法相结合提出将改进的混沌振子方法用于检测低信噪比的正弦信号，仿真结果证明了该方法在微弱信号的探测中的有效性[72]。国育家采用滑模变结构控制法改进霍尔姆斯型 Duffing 系统，仿真结果表明改进后的系统可以有效抑制噪声，并通过系统的功率谱图检测出弱信号的频率[73]。在其他管理科学领域中也开始注重弱信号的研究，衍生出一系列相适应的识别方法。

3.3.1 环境扫描法

弱信号与噪声混杂在一起，在传递过程中，噪声会降低弱信号的可探测性，需要将其过滤掉。但是从模糊的周边环境获取并解读弱信号是非常复杂的过程，过滤不当的话容易得出错误的结论。

H. I. Ansoff 和 E. J. Mcdonnell 提出从不确定性的感知到最后辅助决策，需要经过 3 层筛选：①监视层：弱信号在这层被捕获，从企业所在的商业环境中收集各类信息；②心智层：将获得的弱信号进行迭代，通过经验和假设评估信号的关联性和价值；③推进层：将形成的信息用于决策支撑[53]。R. Wagner 等提出网络环境下的新型扫描方法，运用信息觅食理论、向量空间模

型等逐步过滤,获取有潜在价值的弱信号[74]。

E. Amanatidou 等认为使用地平线扫描弱信号要有两个阶段:①探索式扫描,是指用关键词进行广泛扫描,之后进行评估、选择和聚类;②议题聚焦式扫描,是指对已获得的弱信号进行评价,分析额外的信号来源,进行次级弱信号识别[75]。

3.3.2 情景分析法

情景分析法主要是通过详细、缜密的推理将弱信号放大,建构出未来多种可能的情景。通过不断地信息修正,使有价值的弱信号逐渐明晰,减少认知上的偏差[64]。但该方法所需的推理环节比较多,不断修正情景耗费的时间比较长,同时推理具有较强的主观性,对研究人员的背景知识要求较高。

P. J. H. Schoemaker 等结合情景规划法和商业分析方法等,加快弱信号从弱转强的速度,设计出"战略雷达"工具,交付美国国防后勤局使用[63]。V. Kaufmann 和 E. Ravalet 对样本人口进行调查,重点关注人口移动相关的弱信号,借助情景规划的方法推测出 2050 年法国可能存在的人口流动趋势[76]。P. Meissner 等构建了整合内外部专家判断的结构化框架,用于监测和讨论情景规划过程中的盲点和微弱信号,并以德国建筑业为例,论证了方法的有效性[77]。

3.3.3 模糊综合评价法

模糊综合评价法是一种基于模糊数学的综合评价方法,依据模糊数学的隶属度理论量化不完全信息。处理时的主要步骤有:①信息筛选:根据任务需求收集信息,并对收集到的模糊信息进行筛选;②信息测度:根据研究精度给筛选出的信息评判等级,依据其重要性确定权重;③计算综合评价:采用模糊算子综合评价结果[78]。弱信号携带的信息具有不确定性,该方法可以将对信息的模糊判断转化成定量分析,预测变化的趋势。该方法适用于需要尽快获得分析结果,但信息量较少的情况。

邓胜利等引入模糊集合理论,运用层次分析法确定弱信号特征因素的权重,通过隶属度计算对企业竞争的弱信号进行定量识别[57]。董尹和刘千里在供应链风险弱信号的识别中引入模糊逻辑,运用模糊 TOPSIS 方法进行观测描述,借助仿真算例检验了方法的适用性[79]。这种方法在将模糊信息定量化处理时,权重赋值受主观影响较大,缺乏客观性。

3.3.4 突变理论识别法

突变现象在自然界中是普遍存在的,科学的发展与变革也具有渐变到突变、自组织等特征[80]突变理论是用数学模型来描述连续性行动突然中断导致的质变过程,被用来认识和预测复杂的系统行为[81]。当系统达到某个临界点时,很小的扰动都有可能引起突变。这时系统从扰动中恢复的速率变慢,恢复能力变小的现象称为临界慢化,该现象的前兆信号有:恢复变慢、系统自相关系数增大、波动方差增大[82]。这种识别方法的关键在于检测到微弱前兆信号后采取行动防止系统过渡或为过渡提前做好准备。

吴浩等利用方差和自相关系数系数这两个指标分析计算序列(原始观测资料与参考气候态的距平),检测气候突变前兆信号的可捕捉性,并探讨了两种前兆信号对信噪比不同的资料的适用性[83]。M. Scheffer 认为临界慢化现象仅能说明过渡到新的未知状态的可能性增加,但并非是转换状态的特定预警信号[84]。R. J. Perla 和 J. Carifio 指出库恩提出的渐进性创新和变革性创新分别对应突变现象中原本的连续性变化和在临界点上的突变,所以运用突变论的形式来构架科学认识论的框架是合乎逻辑的[85]。

3.3.5 机器学习手段检测法

大数据时代,弱信号的识别越来越依靠专门的技术手段。当前主要通过文本挖掘、贝叶斯网络、潜在语义分析和局部异常因子等方法对弱信号进行侦测,获取可能存在的弱信号。

J. Yoon 提出一种基于关键字的文本挖掘方法来识别弱信号主题,并利用与太阳能电池相关的网络新闻对该方法进行验证,从可见度、扩散程度和速率增加性几个指标来识别弱信号[86]。D. Thorleuchter 和 D. P. Vanden 运用潜在语义索引和文本挖掘的方法对网络信息进行聚类分析,抽取出有用的弱信号[87-88]韩国科技信息研究院(Korea Advanced Institute of Science and Technology)开发的新兴趋势信号(New and Emerging Signals of Trends, NEST)模型,搜集来自全球专家网络的信息,结合定性和定量方法,使用贝叶斯网络聚类,通过模式识别和交叉影响分析系统地探测弱信号[89]。J. Kim 和 C. Lee 应用文本挖掘、局部利群因子及信号组合图等组合方法提高检测微弱信号的灵敏度[90]。G. Joanny 等运用文本挖掘技术从文档语料库中生成多词概念词典,然后获取包含词典内主题词的数据集,经过 3 层过滤器后获得最终符合突破性技术预测的弱信号数据集[91]。

3.3.6 弱信号识别方法的比较分析

捕获到的弱信号的质量与分析效果不仅与早期信号的质量和数量有关,还取决于研究人员的背景知识、

经验等认知水平。上述不同的识别方法适用于不同的情况,优、缺点也各不相同,如表6所示:

表6 弱信号识别方法对比分析

方法	领域/范围	优点	不足
环境扫描法	企业战略规划	①扫描获取的信息更加丰富;②筛选器可以层层过滤掉噪音	对企业情报分析人员的专业能力要求较高
情景分析法	宏观层面(国家或全行业)	①将弱信号放大;②假设多种不同场景并不断补充信息,减少认知偏差	①主观性较强;②不断补充信息进行修正耗时较长
模糊综合评价法	企业竞争情报	①将模糊的定性信息转化为定量信息;②可用于快速决策	转化时的标准及计算时的各项赋值主观性强
突变理论识别法	自然灾害及自然科学	在临界阈值附近探测预警信号,有效避免系统突变或为突变提前做好准备	临界慢化现象意味着突变的可能性较大,但不一定发生,据此识别弱信号的准确度不高
机器学习手段检测法	大数据时代(爆炸的网络信息)	①可处理的数据量较大;②识别出的弱信号质量较高,可用于精准决策	①对研究人员的专业能力要求较高;②大体量数据处理耗时耗力

由此可见,弱信号的识别过程是一个伴随不确定的复杂过程,需要收集尽可能多的不完全信息,从中遴选出与研究对象相关性较强的弱信号并加以利用。突破性创新在萌芽之初也是以主题间的微弱关系的形式存在。弱关系作为一种特殊的弱信号,携带的信息往往更多元化,能反映领域的整体趋势,因此可以提取有效的弱关系,找出它们的内在联系,实现突破性创新的早期识别。

4 弱关系分析及其应用

4.1 弱关系的内涵

弱关系一词起源于社会学,M. S. Grannovetter 指出强连接关系代表着行动者彼此具有高度的互动,传递的信息同质化程度较高。弱关系是不同群体间信息流通的桥梁,具有中介性、异质性的特点,收集的异质性信息能够给企业带来新的机会,有助于技术创新[92]。因此,弱关系是一个与强关系相对立的概念,指的是在网络结构中节点间的联系强度低于某一阈值的关系。如图2所示:核心层内各节点(黑色圆点)直接连接,节点间联系较为紧密,相对来说属于强关联关系,类似于知识网络中关键词直接共现。逐渐向外扩散到中间层,各节点之间(三角形)关系强度逐渐减弱,外圈层内的节点(正方形)与中心节点需要通过许多中间实体才能产生联系,同时它们彼此之间的联系较少,网络结构较为稀疏。

因为关系的强弱是相对的,具体测度时常按照节点间连线的权重从大到小进行排序,设定合适的阈值,小于该阈值的认定为弱关系[93]。阈值确定也是弱关系分析中的关键一环,需要参考数据集的特点和研究目的经过多次调整试验来确定。

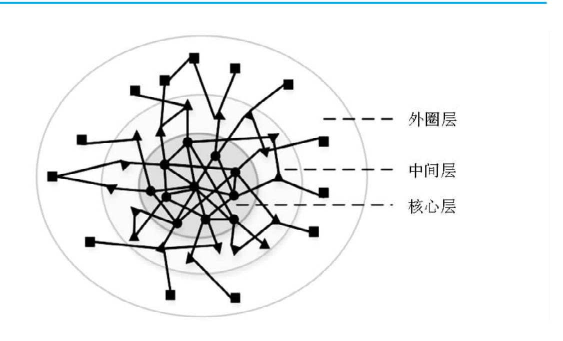


图2 网络结构中的弱关系

4.2 弱关系分析的应用

弱关系理论早期是从社会学的角度研究社交网络的特点,逐渐被应用到企业创新、经济管理以及科技信息等领域。P. A. Julien 等在探讨中小型陆基物流公司的技术变革时指出弱关系比强关系更能促发技术革新,对各类弱关系的分析利用能力越强,公司变革也越快[94]。J. Zenou 等调研工人通过强关系和弱关系的网络获得工作机会的情况,结果表明维系弱关系可以有效提高工人就业率,加快经济发展[95]。刘俊婉等通过构建美国科学院生物学领域244位院士的科学合作网络,对合作的弱关系、强关系以及超级关系特征进行量化分析,发现与名人科学家合作由弱关系成为强关系会加速科研事业的上升[96]。S. H. Yoo 和 D. K. Won 认为应该重点关注未被选为文章关键词的词之间的弱关系,运用基于代理的模拟工具 NetLogo 探索纳米材料领域的技术创新动态行为,识别该领域的未来技术趋势[97]。

单一类型的关系只能反映研究领域的局部特征,因此要将弱关系分析应用于突破性研究主题的早期识别中,仅依靠某一种弱关系是不够的。多元关系融合算法模型可以提供方法上的借鉴,使多种弱关系的有

效融合成为可能,获得更为明确的信息指向。

4.3 弱关系的融合汇聚分析

该方法是将获取到的不同类型的弱关联关系运用融合算法进行综合分析,使弱关系逐级增强,携带的信息逐渐明晰,达到识别突破性创新的目的。由图3可见:最左边一列是各种类型的弱关系,阶段一、阶段二、阶段三是弱关系的融合汇聚过程。首先汇聚同类型的弱关系,使微弱关系变得明显。然后运用算法模型挖掘各类弱关系背后的关联,人为地进行增强,获取背后的信息指向。最后融合多种增强后的弱关系,试图找到突破性研究主题。整个处理过程会过滤掉那些汇聚以后也没有任何价值的弱关系,使结果更精准。

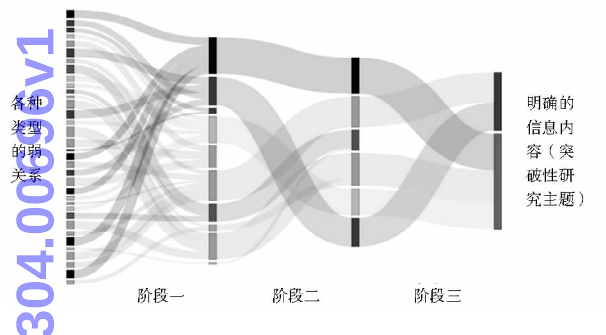


图3 多种弱关系的融合汇聚

当前已有不少学者对多元关系的融合进行了探索。许海云等以基因工程疫苗研发与制备领域为例,利用PathSelClus算法实现共现关系、引文关系与合著关系的融合,对比分析证明,多元关系融合可以有效提高主题聚类的效果[98] S. Jensen等采用元路径方法将文献、主题词、作者、引文等属性关联起来,从不同角度研究对象之间的关联性和相似性,并用于探寻主题演化[99]。D. Zhang等针对元路径严格按照设定选取节点的不足,设计基于元图的方法引导随机游走过程,捕获更为复杂的节点关系[100]。刘彤等利用元矩阵对由专利及其属性构成的关系网络进行表征,引入LDA算法模型融合多重关系,更准确地识别核心专利[101]。C. Shi等提出HERec模型抽取不同类型的节点序列进行表示学习,利用融合函数对表示结果进行融合,实现社区的识别预测[102]。现有方法已经开始注意到分析对象及其关系的复杂性,未来研究需要构建有效的算法模型对多个对象或对象之间多种弱关系反映出来的各种信息进行充分融合。

4.4 高阶网络中的复杂依赖关系

现实世界中的网络环境越来越复杂,传统网络模型难以捕获到节点间更为复杂的依赖关系及路径信息,尤其是在复杂网络中某一节点常常通过若干中间实体与另一节点产生的关联关系,这种弱关联关系在数量上远多于强关系。借鉴高阶网络模型可以捕获分析弱关系,使得模糊信息通过增强、放大分析变得明晰,降低科学发展的不确定性。目前对于高阶网络的研究主要有3个方向:多层高阶模型、组合高阶模型和非马尔可夫高阶模型。

4.4.1 多层高阶模型

该模型针对网络系统实体间多种链接类型的复杂关系。多层高阶模型展现了实际的复杂网络中多种关系类型的交互作用,马蒙周为了提升生物体中度小关键基因的识别效率,首先将同源数据集成Yeast和Human单层网络构建多层复杂网络,并通过图论的方式表示;然后定义概率转移矩阵描述单位时间内基因节点在层内和层间游走动态的转换规则,结合随机游走算法模型选取合适的游走系数,将随机游走后节点的停留概率作为多层网络的中心性指标进行排序,提取出关键基因;最后验证该方法有明显的优越性:识别出的关键基因更多更准确[103]。为了最大限度的保有网络细节信息,有学者引入"张量"对多层网络进行分析。L. Gauvin等将多层网络表示成三阶张量,展示出三个维度间的对应关系。并对网络进行非负张量分解,分解产生的二维矩阵除了包含2个节点与社区间的隶属关系,还将得到1个社区对应的第三个维度的关系信息,如时序网络中可以得到对应时间维度上的活动模式的二维矩阵,该方法可以充分利用各个维度的信息,从而实现社区识别[104]。

4.4.2 组合高阶模型

该模型将代表了复杂系统中实体间的相互影响的成对链接推广到任意节点集的广义链接和路径。组合高阶模型已经从拓扑结构分析推广到了图的高位分析,罗永恩等基于超图模型,结合共享熵构建出新的特征整合方法,以每个特征为顶点、特征之间的多元关系作为超边,应用超图分割算法实现原始复杂网络的聚类划分,该方法识别准确率高于线性支持向量机、多核学习等融合方法[105]。A. R. Benson等为探索复杂网络中的节点连接间的丰富信息,构建了基于高阶连通模式的聚类网络通用框架,将主题分析和网络分区统一起来,揭示复杂系统中的新组织模式和模块[106]。

4.4.3 非马尔可夫高阶模型

该模型引入时间序列数据,注重挖掘依赖关系的真实连接状态,揭示了网络节点之间的间接影响路径。J. Xu 等指出当分析来自复杂系统的序列数据如全球航运流量时,一阶马尔科夫模型认为网络的下一个状态转变仅依赖于当前节点的状态,忽略了该转变可能是前几个转变共同导致的结果,使得预测产生偏差,研究反映了对非马尔可夫模型的需求[107]。R Lambiotte 等对比分析了在社区检测、节点排序和系统动态演化 3 个例子中传统的网络与非马尔可夫高阶模型分析预测的结果,证明非马尔科夫模型能够更好地发掘数据背后的内在联系,反映复杂网络动态变化的真实情况[108]。

5 弱信号识别突破性研究的发展趋势

5.1 捕捉有效的突破性创新早期弱信号

突破性创新研究在萌芽初期往往是以主题之间相互连接的微弱关联的形式存在,所以弱信号是探索突破性创新的一个很好的切入点。不同于高频词或高引用文献此类的强信号,弱信号蕴含的信息更加多元化,能够反映学科的整体知识结构。但现实中很多科学发现和突破性创新,可供人们提前监测捕捉的早期迹象较少。而且研究获取到的弱信号常混杂着噪音,携带的往往都是零碎的不完全信息,能否对其进行有效判别不仅取决于研究人员的知识储备和经验累积,还取决于先进技术的支撑。未来研究中可以多注重捕捉有效弱信号,通过提高弱信号的质量,提高识别结果的准确性。

5.2 重视突破性研究主题的动态演化过程

突破性创新及创造性的发现不是凭空产生的,而是在渐进性创新的基础上发展而来,由量变到质变,实现突破。现有的识别方法大都只关注某个时间点下的静态信号,不能很好地观测整个学科领域知识扩散的状况。明晰学科知识的演变发展路径,结合主题与主题之间的关联关系,可以动态地观察主题产生和发展的演化规律。了解学科领域内的主题发展状况之后,才有可能捕捉到最具突破性变革潜力的主题,进而跟踪监测,发现更多早期征兆。因此,未来研究可以考虑引入时序分析方法,剖析主题的产生、发展直至消亡的规律,尽早识别突破性创新。

5.3 利用高阶网络模型分析弱关系

在复杂网络中引入高阶网络模型能够关注更多维度的信息:①模型不只关注节点之间的相互影响,还关注到网络中边与边的相互影响;②将节点之间的连接关系分为不同的类型;③考虑节点之间的关系对全局(网络中每个点)的影响;④关注到节点间关联关系发生的时间点和先后顺序。以上这些大部分都是运用传统网络分析方法时难以捕获到的信息——弱关系,弱关系在知识网络图谱中对于知识结构的形成发挥了重要的作用,但是背后的信息不太容易被感知。利用高阶网络模型可以更好地挖掘相互影响的节点间真实的依赖关系,实现弱关系的有效融合。所以未来研究可以尝试运用高阶模型,帮助捕获网络中节点之间复杂的弱关联依赖关系,探寻真实的路径信息,从中发现知识发展动向,追踪主题的动态演化过程,实现突破性研究的早期识别。

6 结语

本文首先对比分析了现有的突破性创新主题识别方法,发现当前研究对弱信号的关注不够,同时缺乏对学科主题动态演化过程的深入探索,难以实现早期识别。之后,介绍了弱信号的内涵、特征,归纳总结了几种弱信号的捕获分析方法,阐明弱信号分析用于突破性创新早期识别的可行性。考虑到突破性研究主题的识别预测大多关注文献间的引用关系、主题词间的语义关系等关系类型数据,接下来研究着重分析了弱信号的特殊类型——弱关系,单一弱关系反映的信息有限,借鉴融合算法模型能够实现多种弱关系的融合汇聚,在尽可能早的阶段获得明确的指向性信息。最后,提出未来研究可以通过精准捕捉突破性创新早期的有效弱信号,分析突破性创新主题的动态演化规律,如利用高阶网络模型捕获分析弱信号这种途径提高突破性研究早期识别的准确性。

参考文献:

[1] 万宁.浅析颠覆性创新、破坏性创新和突破性创新三者关系[J].商,2015(30):122-123.
[2] DOSI G, NELSON R R. Technological paradigms and technological trajectories[J]. Research policy, 1982, 11(3): 147-162.
[3] TUSHMAN M L, ANDERSON P. Technological discontinuities and organizational environments[J]. Administrative science quarterly, 1986, 31(3): 439-465.
[4] KOTELNIKOV V. Radical innovation versus incremental innova-

tion [M]. Boston :Harvard Business School Press, 2000:41 - 85.
[5] 肖海林，董慈慈. 突破性技术创新研究:现状与展望——基于SSCI和CSSCI期刊的文献计量分析[J]. 经济管理,2020,42(2):192 - 208.
[6] ZHOU K Z, YIM C K,TSE D K. The effects of strategic orientations on technology- and market-based breakthrough innovations [J]. Journal of marketing, 2005, 69(2):42 - 60.
[7] 孙晓雅，陈娟娟. 创新网络关系强度与创新模式关系的研究综述 [J]. 技术与创新管理, 2016, 37(2): 134 - 140.
[8] FORES B, CAMISON C. Does incremental and radical innovation performance depend on different types of knowledge accumulation capabilities and organizational size [J]. Journal of business research,2016,69(2):831 - 848.
[9] 张金柱，张晓林. 基于科技资源的突破性创新指标及识别方法综述 [J]. 图书情报工作, 2012, 56(22): 56 - 61.
[10] KUHN T S. The structure of scientific revolutions[J]. Physics today, 1962, 16(4):69.
[11] WRAY K B. Kuhn and the discovery of paradigms[J]. Philosophy of the social sciences, 2011, 41(3):380 - 397.
[12] ANDERSEN H, BARKER P, CHEN X. The cognitive structure of scientific revolutions [M]. Cambridge: Cambridge University Press, 2006:6 - 10.
[13] 杜建,孙轶楠,张阳,等. 变革性研究的科学计量学特征与早期识别方法[J]. 中国科学基金,2019,33(1):88 - 98.
[14] GALISON P. Refections on image and logic: a material culture of microphysics[J]. Perspectives on science, 1999, 7(2):255 - 284.
[15] VAN RAAN A F J. On growth, ageing, and fractal differentiation of science[J]. Sentometrics, 2000, 47(2):347 - 362.
[16] HOLLINGSWORTH J R. Scientific discoveries: an institutionalist and path-dependent perspective [J]. Biomedical and health research, 2008,72:317 - 353.
[17] 蒋军锋，李孝兵，殷婷婷，等. 突破性技术创新的形成:述评与未来研究[J]. 研究与发展管理, 2017(6): 109 - 120.
[18] 李政，刘春平，罗晖. 浅析颠覆性技术的内涵与培育——重视颠覆性技术背后的基础科学研究[J]. 全球科技经济瞭望, 2016, 31(10): 53 - 61.
[19] 李良德，陈劲，莫昕玮. 突破性创新管理模式研究[J]. 电子政务, 2001 (11): 38 - 41.
[20] 梁正，邓兴华，洪一晨. 从变革性研究到变革性创新:概念演变与政策启示[J]. 科学与社会, 2017 (3): 94 - 106.
[21] WINNINK J. Early-stage detection of breakthrough-class scientific research [D]. Leiden: Universitite Leiden, 2017.
[22] 付玉秀，张洪石. 突破性创新:概念界定与比较[J]. 数量经济技术经济研究, 2004, 21(3): 73 - 83.
[23] KOSHLAND D E. The cha-cha-cha theory of scientific discovery [J]. Science, 2007, 317(5839):761 - 762.
[24] CAMPANARIO J M. Rejecting and resisting Nobel class discoveries: accounts by Nobel Laureates[J]. Entometrics, 2009, 81(2): 549 - 565.
[25] Deep Tech.《麻省理工科技评论》2020年"全球十大突破性技术"[J]. 科技中国,2020(3):5 - 11.
[26] 中国科学院. 2019科学发展报告[D]. 北京:科学出版社,2020.
[27] SMALLH G. A co-citation model of a scientific specialty: a longitudinal study of collagen research[J]. Social studies of science, 1977, 7(2):139 - 166.
[28] SCHNEIDER J W, COSTAS R. Identifying potential "breakthrough" publications using refined citation analyses: three related explorative approaches[J]. Journal of the Association for Information Science and Technology, 2017, 68(3): 709 - 723.
[29] WANG J, VEUGELERS R, STEPHAN P. Bias against novelty in science: a cautionary tale for users of bibliometric indicators[J]. Research policy, 2017, 46(8): 1416 - 1436.
[30] 李勇，安新颖，赵迎光，等. 结合知识组织体系的突发主题监测研究[J]. 情报理论与实践, 2013, 36(5):120 - 123,128.
[31] KLEINBERG J. Bursty and hierarchical structure in streams[J]. Data mining & knowledge discovery, 2003, 7(4): 373 - 397.
[32] 张金柱，张晓林. 基于被引科学知识主题突变的突破性创新识别[J]. 数据分析与知识发现, 2016, 32(7 - 8):42 - 50.
[33] YOON J, KIM K. Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks[J]. Scientometrics, 2011, 88(1): 213 - 228.
[34] 罗素平，寇翠翠，金金，等. 基于离群专利的颠覆性技术预测——以中药专利为例[J]. 情报理论与实践,2019,42(7): 165 - 170.
[35] PALOMERAS N. Sleeping patents: any reason to wake up? [J]. Iese research papers, 2003, 20(35): 1 - 25.
[36] 杜建."睡美人"文献的识别方法与唤醒机制研究[D]. 南京:南京大学,2017.
[37] VAN RAAN A F J, WINNINK J J. Do younger sleeping beauties prefer a technological prince? [J]. Scientometrics, 2018, 114 (2): 701 - 717.
[38] SOOD A, TELLIS G J. Technological evolution and radical innovation[J]. Journal of marketing, 2005, 69(3): 152 - 168.
[39] TELLIS G J. Disruptive technology or visionary leadership? [J]. Journal of product innovation management, 2006, 23(1): 34 - 38.
[40] 杨国忠,陈佳. 企业突破性技术创新行为研究——基于前景理论的演化博弈分析[J]. 工业技术经济,2020,39(5):57 - 64.
[41] MAO J, MA C, LIANG Z. Identifying emerging technology:a neural network based solution[EB/OL]. [2020 - 03 - 23]. https://vpinstitute. org/wp-content/uploads/2019/10/MTEGTM2019-Jin.pdf.
[42] 徐路路,王芳. 基于支持向量机和改进粒子群算法的科学前沿

预测模型研究[J].情报科学,2019,37(8):22-28.
[43] 刘博文,白如江,周彦廷,等.基金项目数据和论文数据融合视角下科学研究前沿主题识别——以碳纳米管领域为例[J].数据分析与知识发现,2019,3(8):114-122.
[44] 张丽华.研究前沿探测及演化分析方法与实证研究[D].北京:中国科学院大学,2015.
[45] 许海云,董坤,隗玲.学科交叉主题识别与预测研究[M].北京:科学技术文献出版社,2019:53-54.
[46] 罗瑞,许海云,董坤.领域前沿识别方法综述[J].图书情报工作,2018,62(23):119-131.
[47] 叶向东.关于科学不确定性的若干思考[J].全球科技经济瞭望,2008,23(2):32-34.
[48] 张华.论不完全信息条件下的片断情报分析法[J].情报理论与实践,2008(4):498-501.
[49] 莫乃兴.产品信息不对称下的市场信号[J].商场现代化,2007(13):38-39.
[50] HEIL O,ROBERTSON T S. Toward a theory of competitive market signaling: a research agenda[J]. Strategic management journal, 1991,12(6):403-418.
[51] DAY G S,SCHOEMAKER PAUL J H. Peripheral vision: detecting the weak signals that will make or break your company[M]. Boston:Harvard Business School Press,2006:4-6.
[52] ANSOFF H I. Managing strategic surprise by response to weak signals[J]. California management review,1975,18(2):21-33.
[53] ANSOFF H I,MCDONNELL E J. Implanting strategic management [M]. Upper Saddle River:Prentice Hall International Inc,1984.
[54] COFFMAN B. Weak signal research, part I: introduction [EB/OL]. [2020-06-20]. http://legacy.mgtaylor.com/mgtaylor/jotm/winter97/jotmwi97.htm.
[55] BROWN J S. Minding and mining the periphery[J]. Long range planning, 2004, 37(2):143-151.
[56] 赵小康.弱信号:识别、探测与应对[J].情报杂志,2010,29(1):159-163.
[57] 邓胜利,林艳青,王野.企业竞争弱信号的特征提取与定量识别研究[J].图书情报工作,2016,60(10):67-75.
[58] KUUSI O,HILTUNEN E,LINTURI H. Heikot tulevaisuussignaalit: Delfoi-tutkimus[J]. Futura, 2000(2):78-92.
[59] MOIJANEN M. Heikot signaalit tulevaisuudentutkimuksessa[J]. Futura,2003(4):38-60..
[60] PITKÄNEN R. Tulevaisuuden tutkimuksesta kilpailuetua-Opponentti [J]. Yritystalous, 2006: 1-2.
[61] HILTUNEN E. The future sign and its three dimensions[J]. Futures, 2008, 40(3):247-260.
[62] LESCA H, LESCA N. Strategic decisions and weak signals[M]. London:ISTE Ltd,2014:17.
[63] SCHOEMAKER P J H,DAY G S,SNYDER S A. Integrating organizational networks, weak signals, strategic radars and scenario planning[J]. Technological forecasting & social change,2013,80(4): 815-824.
[64] 单彬.认知视角下的弱信号分析及实证研究[D].北京:中国人民解放军军事医学科学院,2014.
[65] HILTUNEN E. Where do future-oriented people find weak signal? [M]. Turku:Finland Futures Research Centre,2007.
[66] 党倩娜.新兴技术弱信号监测机制研究[M].上海:上海科学技术文献出版社,2018:26-27.
[67] ROSSEL P. Weak signals as a flexible framing space for enhanced management and decision-making[J]. Technology analysis & strategic management,2009,21(3): 307-320.
[68] PARK C, CHO S. Future sign detection in smart grids through text mining[J]. Energy procedia, 2017, 128: 79-85.
[69] SAUL P. Seeing the future in weak signals[J]. Journal of futures studies,2006,10(3):93-102.
[70] KUOSA T. Different approaches of pattern management and strategic intelligence [J]. Technological forecasting & social change, 2011,78(3):458-467.
[71] 董尹,刘千里,宋继伟,等.弱信号研究综述:概念、方法和工具[J].情报理论与实践,2018,41(10):147-154.
[72] DENG C, ZHANG S. The weak signal detection based on chaos and genetic algorithms[C]//2009 Second International Symposium on Information Science and Engineering. Washington:IEEE Computer Society, 2009: 579-582.
[73] 国育家. 基于混沌理论的弱信号检测方法的研究[D]. 西安:西安科技大学,2017.
[74] WAGNER R,SCHOLZ S W,DECKER R. An internet-based approach to environmental scanning in marketing planning[J]. Marketing intelligence & planning,2005,23(2):189-199.
[75] AMANATIDOU E,BUTTER M,CARABIAS V,et al. On concepts and methods in horizon scanning:lessons from initiating policy dialogues on emerging issue[J]. Science & public policy,2012,39(2):208-221.
[76] KAUFMANN V, RAVALET E. From weak signals to mobility scenarios: a prospective study of france in 2050[J]. Transportation research procedia, 2016, 19:18-32.
[77] MEISSNER P, BRANDS C, WULF T. Quantifiying blind spots and weak signals in executive judgment: a structured integration of expert judgment into the scenario development process[J]. International journal of forecasting, 2017, 33(1): 244-253.
[78] 郑圆圆,陈再良.模糊理论的应用与研究[J].苏州大学学报(工科版),2011,31(1):52-58.
[79] 董尹,刘千里.供应链风险识别中的弱信号介入、感知机制与观测方法研究[J].情报工程,2019,5(3):49-64.
[80] NOYONS E C M, RAAN A F J V. Monitoring scientific developments from a dynamic perspective: self-organized structuring to map neural network research[J]. Journal of the American Society

for Information Science, 1998, 49(1):68 – 81.
[81] 赵松年. 突变理论:形成、发展与应用[J]. 世界科学, 1989(4):12 – 14.
[82] SCHEFFER M, BASCOMPTE J, BROCK W A, et al. Early-warning signals for critical transitions[J]. Nature, 2009, 461(7260): 53 – 59.
[83] 吴浩, 侯威, 颜鹏程. 试用临界慢化原理探讨气候突变[J]. 物理学报, 2016(3):556 – 565.
[84] SCHEFFER M. Foreseeing tipping points[J]. Nature, 2010,467(7314):411 – 412.
[85] PERLA R J, CARIFIO J. The nature of scientific revolutions from the vantage point of chaos theory[J]. Science & education, 2005, 14(3/5):263 – 290.
[86] YOON J. Detecting weak signals for long-term business opportunities using text mining of Web news[J]. Expert systems with applications, 2012, 39(16):12543 – 12550.
[87] THORLEUCHTER D, VAN DEN POEL D. Weak signal identification with semantic web mining[J]. Expert systems with applications, 2013, 40(12): 4978 – 4985.
[88] THORLEUCHTER D, VAN DEN POEL D. Idea mining for web-based weak signal detection[J]. Futures, 2015, 66(Feb.): 25 – 34.
[89] KIM S, KIM Y E, BAE K J, et al. NEST: a quantitative model for detecting emerging trends using a global monitoring expert network and Bayesian network[J]. Futures, 2013, 52(Aug.): 59 – 73.
[90] KIM J, LEE C. Novelty-focused weak signal detection in futuristic data: assessing the rarity and paradigm unrelatedness of signals[J]. Technological forecasting and social change, 2017, 120(Jul.): 59 – 76.
[91] JOANNY G, PERANI S, EULAERTS O. Detection of disruptive technologies by automated identification of weak signals in technology development[C]//Proceedings of the International Conference on Scientometrics and Informetrics. Rome:INT SOC Scientometrics & Informetrics-ISSI,2019: 2644 – 2645.
[92] GRANOVETTER M S. The strength of weak ties: a network theory revisited[J]. Sociological theory,1983,1(1):201 – 233.
[93] WEI L, XU H, WANG Z, et al. Topic detection based on weak tie analysis: a case study of LIS research[J]. Journal of data and information science, 2017, 1(4): 81 – 101.
[94] JULIEN P A, ANDRIAMBELOSON E, RAMANGALAHY C. Networks, weak signals and technological innovations among SMEs in the land-based transportation equipment sector[J]. Entrepreneurship & regional development, 2004, 16(4): 251 – 269.
[95] ZHAO J, WU J, XU K. Weak ties: subtle role of information diffusion in online social networks [J]. Physical review E, 2010, 82(1): 016105.
[96] 刘俊婉,丁凯悦,王菲菲,等.科学合作的弱关系、强关系以及超级关系研究[J].科学学研究,2017,35(4):500 – 510,543.
[97] YOO S H, WON D K. Simulation of weak signals of nanotechnology innovation in complex system[J]. Sustainability, 2018, 10(2): 486 – 499.
[98] 许海云,武华维,罗瑞,等.基于多元关系融合的科技文本主题识别方法研究[J].中国图书馆学报,2019,45(1):82 – 94.
[99] JENSEN S, LIU X, YU Y, et al. Generation of topic evolution trees from heterogeneous bibliographic networks[J]. Journal of informetrics, 2016, 10(2):606 – 621.
[100] ZHANG D, YIN J, ZHU X, et al. Metagraph2vec: complex semantic path augmented heterogeneous network embedding[C]// Pacific-Asia conference on knowledge discovery and data mining. Melbourne:Springer,2018: 196 – 208.
[101] 刘彤, 杨冠灿, 侯元元. 基于多重关系整合的专利网络分析方法研究与应用[J]. 情报理论与实践, 2016, 39(2):59 – 63.
[102] SHI C, HU B, ZHAO X,et al. Heterogeneous information network embedding for recommendation[J]. IEEE transactions on knowledge and data engineering,2019,31(2):357 – 370.
[103] 马蒙周. 基于多层网络的关键基因识别方法研究[D]. 西安:西安理工大学,2018.
[104] GAUVIN L, PANISSON A, CATTUTO C. Detecting the community structure and activity patterns of temporal networks: a non-negative tensor factorization approach[J]. PloS one, 2014, 9(1): e86028.
[105] 罗永恩, 胡继承, 徐茜. 基于超图的多模态关联特征处理方法[J]. 计算机工程, 2017,43(1):226 – 230.
[106] BENSON A R, GLEICH D F, LESKOVEC J. Higher-order organization of complex networks[J]. Science, 2016, 353(6295): 163 – 166.
[107] XU J, WICKRAMARATHNE T L, CHAWLA N V. Representing higher-order dependencies in networks[J]. Science advances, 2016, 2(5):e1600028.
[108] LAMBIOTTE R, ROSVALL M, SCHOLTES I. From networks to optimal higher-order models of complex systems[J]. Nature physics, 2019, 15(4): 313 – 320.

作者贡献说明:
刘亚辉:负责资料收集整理,论文撰写与修改;
许海云:指导论文撰写与修改。

刘亚辉，许海云.突破性创新早期识别与弱信号分析综述[J].图书情报工作,2021,65(4):89－101.

A Review of Early Recognition of Breakthrough Innovations and the Weak Signal Analysis

Liu Yahui[1,2] Xu Haiyun[3,4]

[1] Chengdu Library of Chinese Academy of Sciences, Chengdu 610041

[2] Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190

[3] Business School, Shandong University of Technology, Zibo 255000

[4] Institute of Scientific and Technical Information of China, Beijing 100038

Abstract: [**Purpose/significance**] This paper has summarized disadvantages of various recognition methods of breakthrough researches by comparing and analyzing these methods. Then this study has introduced weak signals into the identification of radical innovation, focusing on various types of weak signal (especially weak ties) with a view to achieving early prediction. [**Method/process**] Firstly, by analyzing the existing recognition methods, this paper summarized the main problems and pointed out the necessity of weak signal research. Then, this study introduced the concept and representation of weak signal from different disciplines summarized the characteristics of weak signal and combed several methods of weak signal identification. Finally, this paper introduced the connotation and application of weak ties, and proposed that the algorithm model that integrates multiple relationships can realize the effective fusion of various weak ties and obtain more clear directional information. [**Result/conclusion**] In the frontier research, most attention is paid to strong ties such as reference relations between literatures and semantic relations between topic terms. However, weak ties contain more diverse information, and strengthening the analysis of them can enable early recognition of breakthrough research frontier. In the future, it is necessary to seek effective methods to capture weak ties and further dig out the evolution of the theme such as: using high-order network models to analyze weak ties to improve the accuracy of early recognition.

Keywords: radical innovation early recognition weak signal weak linkage multi-relations fusion